Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 4976–4980, July 1990
Medical Sciences

Cloning and expression of a membrane antigen of *Entamoeba histolytica* possessing multiple tandem repeats

(amebiasis/pathogenesis/membrane proteins/adhesion)

SAMUEL L. STANLEY, JR.*†, ANNETTE BECKER*, CYNTHIA KUNZ-JENKINS*, LYNNE FOSTER*, AND ELLEN LI*‡

Departments of *Medicine, †Molecular Microbiology, and ‡Biochemistry and Molecular Biophysics, Washington University School of Medicine, Saint Louis, MO 63110

Communicated by Stuart Kornfeld, April 6, 1990 (received for review February 12, 1990)

ABSTRACT ***Entamoeba histolytica*** **causes amebic dysentery and amebic liver abscess, major causes of morbidity and mortality worldwide. We have used differential hybridization screening to isolate an *E. histolytica*-specific cDNA clone. The cDNA was found to encode a serine-rich *E. histolytica* protein (SREHP) containing multiple tandem repeats. The structural motif of SREHP resembles some of the repetitive antigens of malarial species, especially the circumsporozoite proteins. A recombinant *trpE* fusion protein containing the tandem repeats of SREHP was recognized by immune serum from a patient with amebiasis, demonstrating that SREHP is a naturally immunogenic protein. An antiserum raised against the recombinant fusion protein specifically bound to two distinct bands with apparent molecular masses of 46 and 52 kDa in a crude preparation of *E. histolytica* trophozoite membranes. This antiserum also inhibited *E. histolytica* trophozoite adhesion to Chinese hamster ovary cells *in vitro*. The ability to isolate *E. histolytica*-specific genes, and to express those genes in *Escherichia coli*, may be important in studying the molecular basis of *E. histolytica* pathogenesis and for the future development of vaccines.**

The protozoan pathogen *Entamoeba histolytica* is a major cause of debilitating illness and death worldwide, infecting 500 million people and causing an estimated 50 million cases of diarrhea and 50,000 deaths yearly (1). There is an urgent need for a vaccine that could prevent the establishment of *E. histolytica* infection or the development of invasive disease. Studies in animal models have demonstrated that immunity to *E. histolytica* infection can be produced by immunization with *E. histolytica* lysates (2–4). However, the difficulty in obtaining large quantities of trophozoites, and the relatively crude nature of the immunizing preparations, have severely limited the scope of these studies. To approach this problem, we isolated genes expressed in a pathogenic strain of *E. histolytica* but not in the nonpathogenic *E. histolytica*-like Laredo strain. Our objectives were 2-fold: (*i*) to isolate molecules that may play a role in the pathogenesis of amebiasis, and (*ii*) to isolate *E. histolytica*-specific antigens with vaccine potential.

Here we report the isolation and expression in *Escherichia coli* of an *E. histolytica*-specific cDNA clone that appears to encode an antigenic membrane protein possessing multiple tandem repeats.§ Preliminary studies with antiserum to the recombinant fusion protein suggest that this molecule could play a role in amebic adhesion to a mammalian cell line.

MATERIALS AND METHODS

E. histolytica **Isolates and Culture Conditions.** The *E. histolytica* strain HM1:IMSS is virulent both *in vivo* and *in vitro* (5, 6). The *E. histolytica*-like Laredo strain was isolated from a patient with diarrhea (7), but it is avirulent in *in vitro* cytotoxicity assays and in animal models (5, 6). Both strains were grown axenically in TYI-S-33 medium as described (8).

Construction and Screening of a HM1:IMSS cDNA Library. Total cellular RNA was isolated from exponentially growing HM1:IMSS and Laredo trophozoites by the method of Chirgwin *et al.* (9). Poly(A)$^+$ RNA was purified by chromatography on oligo(dT)-cellulose (10) and was used to prepare double-stranded cDNA (11). The double-stranded cDNA was C-tailed by using the enzyme terminal deoxynucleotidyltransferase and annealed with G-tailed *Pst* I-digested pUC13 (12). The chimeric plasmid was used to transform DH-5 *E. coli* (13), generating a cDNA library containing 50,000 recombinants. Replica filters of 10,000 recombinants were probed with radiolabeled single-stranded cDNA from HM1:IMSS or the nonpathogenic, *E. histolytica*-like Laredo strain, respectively.

Northern Blot Analysis. Northern blots were prepared after 20-μg samples of total cellular RNA were subjected to electrophoresis through 1.5% agarose gels containing formaldehyde (14). The Northern blots were probed with cDNA labeled with [α-^{32}P]ATP to a specific activity of 1000 cpm/pg by the random primer method (15). The hybridization and washing conditions were identical to those itemized in ref. 16 except the last three washes were done at 60°C. Autoradiographs were exposed for 24 hr.

Nucleotide Sequence Analysis. The cDNA clones were sequenced by the technique of Maxam and Gilbert (17). The cDNA sequence of clone c1 was completed by primer extension as described (18), utilizing the oligonucleotide TTCAGGACTAGCTTCGTTCTT derived from the sequence of c1.

Construction of *trpE* Hybrid Gene Fusions. The pATH2 and pATH3 plasmids (described in ref. 19) were used to express both open reading frames of the c1 cDNA clone in *E. coli* as *trpE* fusion proteins. The *Hin*dIII/*Sma* I fragment of c1 (containing nucleotides 128–722 and a portion of the pUC13 polylinker region) was ligated into the *Hin*dIII site and the *Cla* I site (which had been made blunt-ended with Klenow DNA polymerase) of pATH3 to construct pORF1. The *Hin*dIII/*Sma* I fragment was ligated into the *Hin*dIII and blunt-ended *Cla* I sites of pATH2 to construct pORF2.

Expression and Partial Purification of the *trpE*-Encoded Fusion Proteins. The *trpE*-encoded fusion proteins were expressed as described by Hardy and Strauss (19) except the bacteria were harvested after incubation with β-indoleacrylic acid for 24 hr. An insoluble protein fraction was prepared from pelleted cells as described (19) to partially purify the

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: ORF, open reading frame; SREHP, serine-rich *Entamoeba histolytica* protein.

§The sequence reported in this paper has been deposited in the GenBank data base (accession no. M34438).

trpE-encoded proteins. The yield of fusion protein as determined by discontinuous SDS/PAGE using molecular mass standards for comparison was 15–30 μg of protein per ml of culture.

Immunizations. Polyclonal rabbit serum directed against HM1:IMSS trophozoites was obtained by immunizing New Zealand White rabbits subcutaneously with a preparation of 2×10^6 HM1:IMSS trophozoites suspended in complete Freund's adjuvant. Booster immunizations were performed with the same preparation in incomplete Freund's adjuvant. Serum was first collected 2 weeks after the initial booster injection and again 2 weeks after each consecutive booster injection.

Gel purification of the fusion proteins, and immunization of female New Zealand White rabbits with fusion proteins, was performed exactly as described (19) using 75–100 μg of protein per injection. Serum was collected from each rabbit before immunization to serve as a control, then 6 weeks after the initial immunization (2 weeks after the first booster injection), and 2 weeks after each consecutive booster injection.

Western Blots. For Western blotting of the fusion proteins with *E. histolytica* immune human and rabbit serum, the insoluble protein fraction of bacterial cells from 0.3-ml cultures was separated by SDS/10% PAGE under reducing conditions and then transferred to nitrocellulose. Blots were reacted with immune or preimmune serum diluted 1:500, and immunoglobulin binding was detected with ^{125}I-labeled staphylococcal protein A. Blots were autoradiographed for 12 hr.

For Western blotting of amebic lysates with antiserum against recombinant proteins, 5×10^6 trophozoites from 72-hr cultures were washed twice in phosphate-buffered saline (PBS), then suspended in 2 ml of PBS containing 5 mM EDTA, 2 mM leupeptin, 5 mM *N*-ethylmaleimide, 2 mM phenylmethylsulfonyl fluoride, 2 mM benzamidine, and 5 mM *trans*-epoxysuccinyl-L-leucylamido(4-guanidino)-butane (E-64). The trophozoites were lysed by sonication, and homogenates of the HM1:IMSS trophozoites were centrifuged at $100,000 \times g$ for 1 hr. Aliquots of the $100,000 \times g$ supernatant and pellet fraction corresponding to 10^5 trophozoites were denatured, separated by SDS/10% PAGE under reducing conditions, and then transferred to nitrocellulose. Blots were reacted with a 1:1000 dilution of rabbit antiserum to the fusion protein; subsequent steps were performed as described above.

Adherence of [^{3}H]Thymidine-Labeled *E. histolytica* Trophozoites to Monolayers of Cell Line 1021 Chinese Hamster Ovary Cells. These studies were performed as described (20). Medium (0.25 ml) with or without the antisera to be tested was added to the monolayers immediately before addition of 0.25 ml of trophozoites suspended at 8×10^5 cells per ml. A 1:10 dilution of each of the antisera tested was used. Duplicate wells were used for each assay. Data are presented as percentage trophozoite adherence seen in the medium control group ± SD.

RESULTS AND DISCUSSION

Isolation of a cDNA Clone from the Pathogenic *E. histolytica* HM1:IMSS Strain. We screened 10,000 recombinants from the pUC13 cDNA library derived from HM1:IMSS mRNA with radiolabeled single-stranded cDNA transcribed from poly(A)$^+$ RNA from HM1:IMSS or the nonpathogenic, *E. histolytica*-like Laredo strain, respectively. Four unique clones that hybridized to HM1:IMSS cDNA but not to Laredo cDNA were isolated. In this report, we describe the characterization of one of the clones, c1.

Analysis of c1 Gene Expression by RNA Blot Hybridization. The expression of the gene corresponding to the c1 clone was examined by blot hybridization studies of RNA isolated from four *E. histolytica* strains (HM1:IMSS, HK9; NIH:200, Rahman), Laredo, *Entamoeba invadens*, and *Entamoeba moshkovskii* (Fig. 1). The c1 clone hybridized with a 0.8-kilobase (kb) species in all four axenic strains of *E. histolytica*, but it did not hybridize with RNA from Laredo, or the two non-*E. histolytica* sp., *E. invadens*, and *E. moshkovskii*. When a Southern blot of *Eco*RI-digested genomic DNA from the same ameba was probed with the c1 clone a 4.4-kb fragment was noted in all four *E. histolytica* strains, but not in Laredo, *E. invadens*, or *E. moshkovskii* (data not shown). The results from RNA and DNA hybridization studies indicate that the gene corresponding to the c1 clone is present and expressed only in the *E. histolytica* strains, and not in the other *Entamoeba* sp. surveyed.

Nucleotide Sequence of the c1 cDNA Clone. The c1 cDNA sequence consisted of 722 nucleotides (Fig. 2*A*) and contained two open reading frames (ORFs), which were designated ORF1 and ORF2. ORF1 contained a continuous ORF from a putative initiator methionine beginning at nucleotide 4 to a TAA termination at nucleotide 703. Beginning at nucleotide 247 ORF1 contained a stretch of 36 nucleotides encoding the dodecapeptide Ser-Ser-Ser-Asp-Lys-Pro-Asp-Asn-Lys-Pro-Glu-Ala. After a 24-nucleotide space encoding a similar octapeptide, the dodecapeptide was tandemly repeated five times, followed by four repeats of an octapeptide, Ser-Ser-Thr-Asn-Lys-Pro-Glu-Ala. The nucleotide sequence of the dodecapeptide repeats was highly conserved, with the only difference being the use of C or U in the third position of the codon for the first serine of each repeat. The repeated octapeptide represents a truncated version of the dodecapeptide, with a single nucleotide change substituting a threonine for the third serine, and nucleotides encoding both aspartic acid residues, and one of the lysine and proline residues absent. Serine constituted 52 of the 233 amino acids; hence, we have referred to the derived amino acid sequence of ORF1 as the serine-rich *E. histolytica* protein (SREHP). Three contiguous serine residues were part of the dodecapeptide repeats and were found in four other locations in the protein. The repeats were preceded by a highly charged region consisting of multiple lysine, glutamate, and aspartate residues. ORF1 terminated with 54 nucleotides encoding primarily hydrophobic amino acids, consistent with a possible membrane insertion or spanning region. The initial 13 amino acids of the N terminus possess some of the characteristics of a eukaryotic signal sequence as defined by the algorithm of von Heinje (22), with a possible cleavage site between the alanine at position 13 and threonine at position 14. The

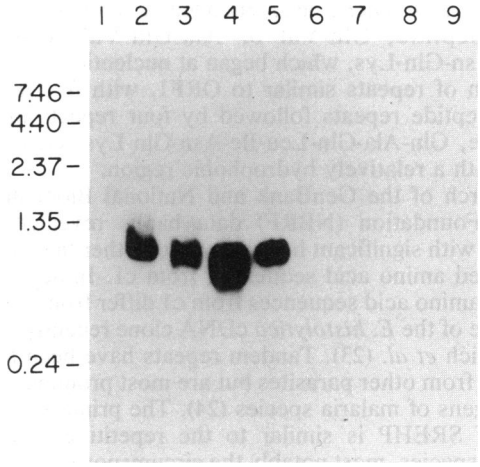

FIG. 1. Northern blot of amebic RNA probed with the c1 clone. Lanes: 1, molecular size standards (kDa); 2, HM1:IMSS; 3, HK9; 4, NIH:200; 5, Rahman; 6, Laredo; 7, *E. invadens*; 8, *E. moshkovskii*; 9, molecular size standards.

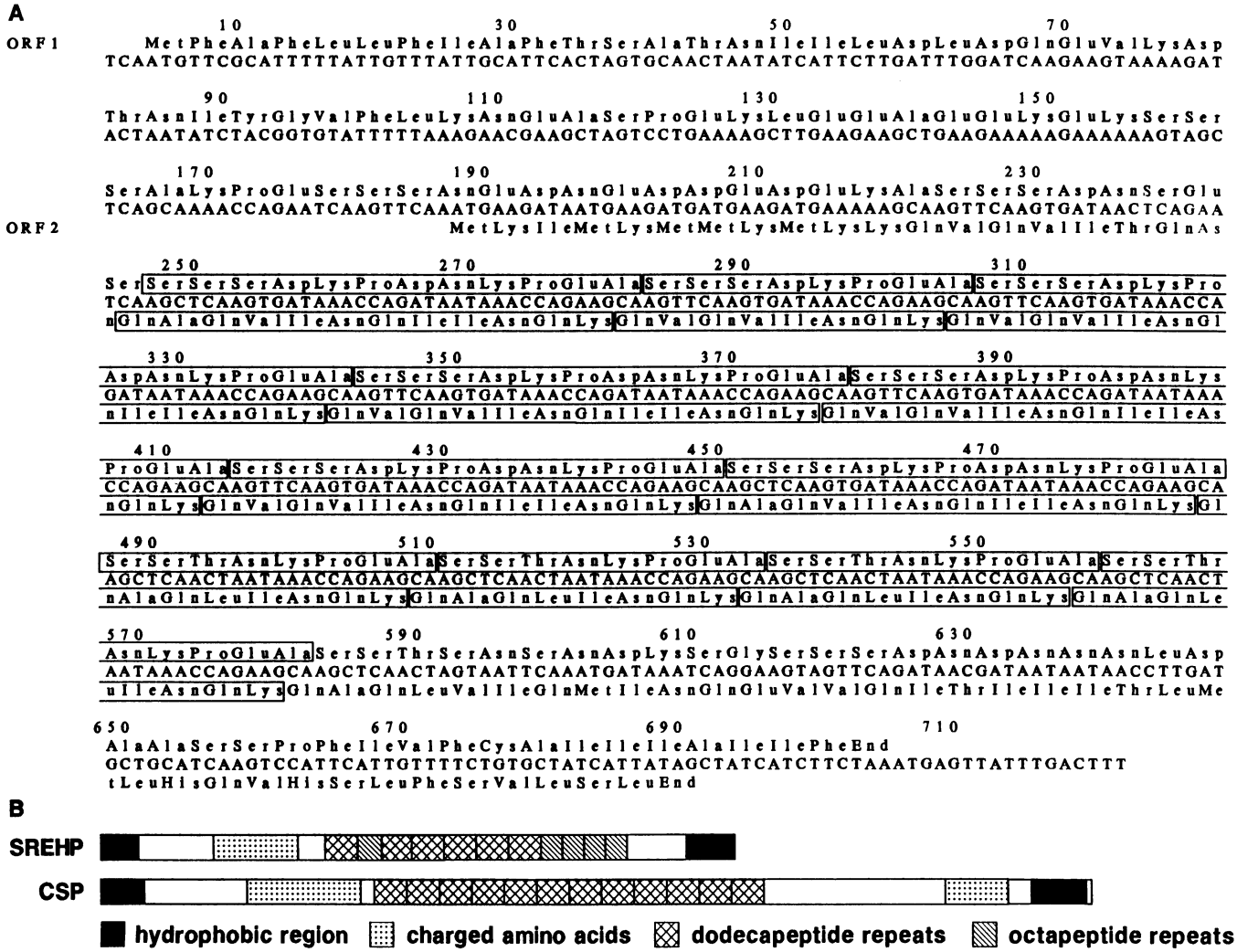

FIG. 2. (*A*) Nucleotide and derived amino acid sequence of the c1 cDNA. The derived amino acid sequence for both ORF1 and ORF2 is shown. The dodecapeptide and octapeptide repeats in both ORFs are boxed. (*B*) General structure of SREHP and the circumsporozoite protein (CSP) of *Plasmodium knowlesi* (21). The *P. knowlesi* dodecapeptide repeat consists of Gly-Gln-Pro-Gln-Ala-Gln-Gly-Asp-Gly-Ala-Asn-Ala (21).

sequence differed from other signal sequences in the absence of one or more charged amino acids (n domain) before the hydrophobic core (h domain).

The second ORF, ORF2, possessed five methionine codons at the 5′ end, with the most 5′ beginning at nucleotide 188 (Fig. 2*A*). ORF2 also encoded a tandemly repeated dodecapeptide, Gln-Val or Ala-Gln-Val-Ile-Asn-Gln-Ile-Ile-Asn-Gln-Lys, which began at nucleotide 245 and had a pattern of repeats similar to ORF1, with five additional dodecapeptide repeats followed by four repeats of an octapeptide, Gln-Ala-Gln-Leu-Ile-Asn-Gln-Lys. ORF2 terminated with a relatively hydrophobic region.

A search of the GenBank and National Biomedical Research Foundation (NBRF) data banks revealed no sequences with significant homology with either the nucleotide or derived amino acid sequences from c1. In addition, the derived amino acid sequences from c1 differ from the partial sequence of the *E. histolytica* cDNA clone recently reported by Tannich *et al.* (23). Tandem repeats have been found in antigens from other parasites but are most prominent among the antigens of malaria species (24). The primary structural motif of SREHP is similar to the repetitive antigens of malarial species, most notably the circumsporozoite proteins (21, 24). Circumsporozoite proteins cover the surface of the sporozoite stage of the malaria parasite, range in size from 40 to 60 kDa, and contain a species-specific pattern of multiple tandem repeats (21, 24). The predicted structure of the *Plasmodium knowlesi* circumsporozoite protein consists of a hydrophobic N-terminal region followed by a series of tandemly repeated amino acids flanked by two domains containing predominantly charged amino acids, and concluding with a C-terminal region, consisting of a hydrophobic anchor region (Fig. 2*B*) (21). The predicted structure of SREHP consists of a hydrophobic N-terminal region, followed by a region containing primarily charged amino acids, followed by tandemly repeated amino acids, and it ends with a C-terminal hydrophobic region (Fig. 2*B*). In the N-terminal charged area of the *P. knowlesi* circumsporozoite protein, 27 of the 48 amino acid residues are charged (primarily lysine and glutamate residues) (21), whereas 21 of the 36 residues are charged (primarily lysine, glutamate, and aspartate) in the comparable N-terminal region of SREHP. Unlike circumsporozoite proteins, SREHP does not possess a second C-terminal group of charged amino acids. An additional similarity between SREHP and circumsporozoite proteins lies in the amino acids used in the repeating units. Circumsporozoite protein repeats appear to be derived from a repertoire of eight amino acids—Ala, Pro, Gly, Asn, Gln, Asp, Arg, and Glu (25). Five of these eight amino acids—Ala, Pro, Asn, Asp, and Glu—are among the seven amino acids that make up the dodecapeptide repeats of SREHP.

c1 cDNA Encodes a Naturally Immunogenic *E. histolytica* Protein. Since we did not know which of the ORFs of c1 were translated in *E. histolytica*, we expressed both ORFs of c1 in

E. coli by using the *trpE*-containing vectors pATH2 and pATH3. We produced constructs pORF1 and pORF2 containing most of the sequence of the c1 insert fused in the appropriate reading frame to the N-terminal two-thirds of the *E. coli trpE* gene. Coomassie blue staining of SDS/polyacrylamide gel-separated insoluble pellets from cells containing the pORF1 or pORF2 vector, which had been induced to produce *trpE* by β-indoleacrylic acid, revealed fusion proteins with molecular masses of ≈60 kDa (the predicted molecular masses for the fusion proteins encoded by ORF1 and ORF2 are 57 and 59 kDa, respectively) (data not shown). To determine whether either of the ORFs of c1 produced a naturally immunogenic *E. histolytica* protein, Western blotting was performed (Fig. 3). Immune serum from a patient with amebic liver abscess and serum from a rabbit immunized with HM1:IMSS trophozoites both bound to a series of bands (largest size, ≈60 kDa) in bacteria expressing the fusion protein (predicted size, 57 kDa) encoded by pORF1 (Fig. 3 *B* and *D*, lane 2). Immune serum did not bind to the fusion protein encoded by pORF2. Control human serum and preimmune rabbit serum did not bind either of the fusion proteins (Fig. 3 *A* and *C*). In subsequent studies, we have found that sera from 25 additional patients with invasive amebiasis also bind the fusion protein produced by pORF1 (S.L.S., A.B., C.K.-J., T. F. H. G. Jackson, S. L. Reed, J. Calderon, and E.L., unpublished data). We have confirmed that the recombinant protein recognized by immune serum contains the tandemly repeated peptides encoded by ORF1 by demonstrating that antiserum to a synthetic peptide containing the dodecapeptide repeat Ser-Ser-Ser-Asp-Lys-Pro-Asp-Asn-Lys-Pro-Glu-Ala binds the recombinant protein (data not shown).

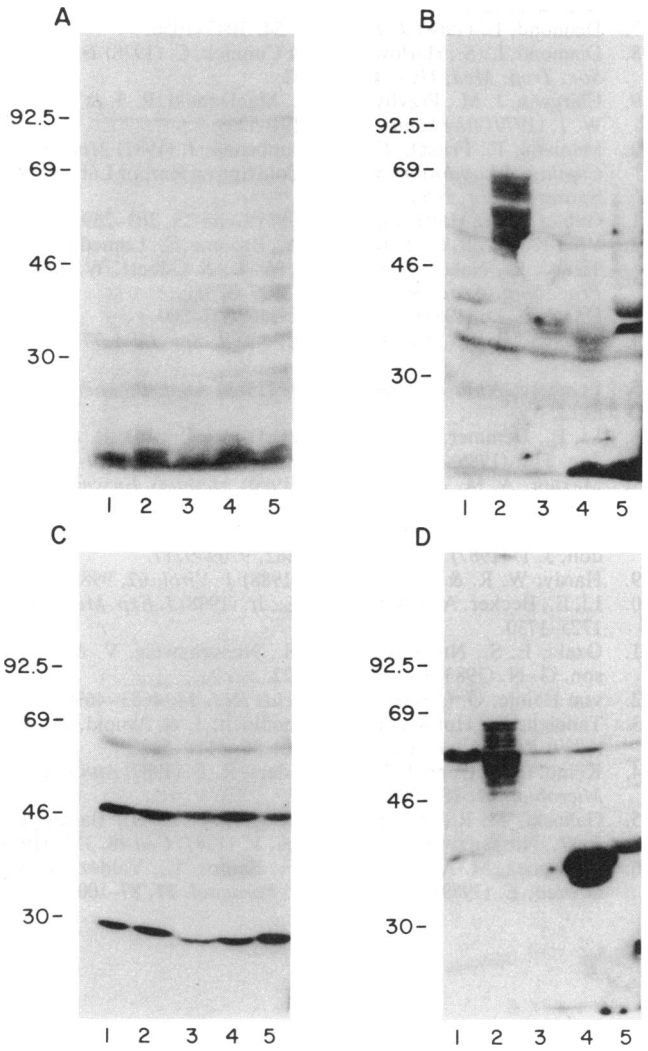

FIG. 3. Western blotting of *trpE*-encoded fusion proteins with normal human serum (*A*), immune serum from a patient with amebic liver abscess (*B*), preimmune rabbit serum (*C*), and serum from a rabbit immunized with HM1:IMSS trophozoites (*D*) was performed as described. The insoluble protein fraction containing the *trpE*-encoded fusion proteins was prepared from bacteria containing the following: lanes 1, pORF2 (encoding the Gln-Val-Gln-Val-Ile-Asn-Gln-Ile-Ile-Asn-Gln-Lys repeats); lanes 2, pORF1 (encoding a portion of SREHP including the Ser-Ser-Ser-Asp-Lys-Pro-Asp-Asn-Lys-Pro-Glu-Ala repeats); lanes 3, pATH2 vector only; lanes 4, pATH3 vector only; lanes 5, lacking plasmid. Molecular size standards in kDa are indicated.

SREHP Is a Membrane Protein. To determine whether SREHP is associated with a membrane fraction of *E. histolytica*, HM1:IMSS trophozoites were lysed, then spun at 100,000 × *g*, and loaded onto SDS/polyacrylamide gel as a supernatant and pellet fraction. As shown in the Western blot (Fig. 4), the anti-SREHP fusion protein antiserum specifically bound to two distinct bands at 46 and 52 kDa in the 100,000 × *g* pellet fraction. Little binding was detected in the 100,000 × *g* supernatant, suggesting that the native SREHP is primarily membrane bound. This finding is consistent with the primary structure data. The anti-SREHP fusion protein antiserum showed no binding to any species in whole Laredo lysates, suggesting that SREHP is *E. histolytica* specific. It is unclear at this time why antiserum to the recombinant SREHP fusion protein detects two species approximately twice the size predicted from the derived amino acid sequence of the SREHP cDNA (25 kDa) in the *E. histolytica* lysates. The results of the Northern blotting study (Fig. 1) demonstrated that the size of the c1 clone (722 nucleotides) is close to the size of the SREHP transcript (≈800 nucleotides), suggesting that most, if not all, of the coding region should be contained in the c1 clone. Hence, it is unlikely that the discrepancy in size results from additional amino acids, and more likely it is secondary to posttranslational modifications. Antiserum to the fusion protein encoded by pORF2 showed no binding to HM1:IMSS trophozoite lysates by Western blotting (data not shown). Thus, we have been unable to demonstrate the existence of a HM1:IMSS trophozoite protein encoded by c1 ORF2. We cannot exclude the

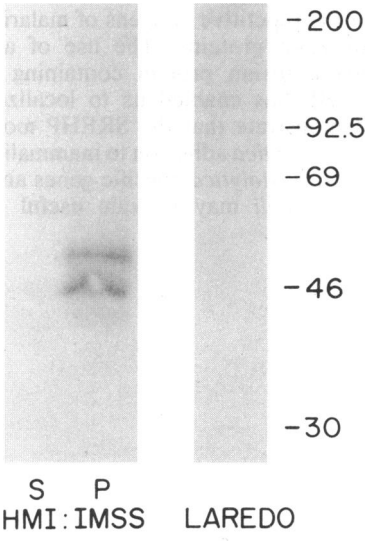

FIG. 4. Identification of the native HM1:IMSS SREHP. Western blot analysis of the supernatant (S) and pellet (P) from the 100,000 × *g* centrifugation of HM1:IMSS trophozoite lysates and whole Laredo lysates by antiserum against the SREHP fusion protein. Molecular size standards in kDa are indicated.

Table 1. Antiserum to the recombinant SREHP fusion protein inhibits the binding of HM1:IMSS trophozoites to the 1021 Chinese hamster ovary cell line

Serum	% of control binding to 1021 cells
Preimmune	94 ± 25
Anti-SREHP fusion protein	30 ± 2
Anti-ORF2 fusion protein	120 ± 25

Results are the means from three experiments and are presented as percentage trophozoite adherence in the medium control group ± SD.

possibility that ORF2 might be translated in the cyst form of the parasite.

Antiserum to the SREHP Fusion Protein Inhibits *E. histolytica* HM1:IMSS Adhesion to Chinese Hamster Ovary Cells. We have previously studied the adherence of *E. histolytica* HM1:IMSS trophozoites to a panel of Chinese hamster ovary cells and shown that trophozoites adhere best to the Chinese hamster ovary cell line 1021 (20). We evaluated the ability of antiserum to the SREHP fusion protein to inhibit the binding of radiolabeled *E. histolytica* HM1:IMSS trophozoites to 1021 cells. We compared antiserum to the SREHP fusion protein with preimmune serum, antiserum to the fusion protein encoded by pORF2, and a control medium. Antiserum to the SREHP fusion protein reduced the binding of HM1:IMSS trophozoites to 1021 cells to 30% ± 2% of control levels (Table 1). Preimmune serum, and antiserum to the fusion protein encoded by pORF2, had no inhibitory effects. This finding suggests that SREHP is located on the trophozoite cell surface and could possibly play a role in *E. histolytica* adhesion. It is notable in this regard that a recent study by Rodriguez *et al.* (26) using a polyclonal antiserum against total amebic proteins found that an *E. histolytica* protein of 50 kDa was one of seven *E. histolytica* proteins found on the surface of erythrocytes incubated with trophozoites.

In summary, our application of cDNA differential hybridization screening using a pathogenic *E. histolytica* strain and a closely related *E. histolytica*-like strain has permitted the isolation of an *E. histolytica*-specific cDNA encoding a serine-rich *E. histolytica* protein, SREHP. SREHP is an immunogenic membrane protein with multiple tandem repeats. The general structure of SREHP resembles the structure of some of the repetitive antigens of malaria, especially the circumsporozoite proteins. The use of an antiserum directed against a fusion protein containing the tandem repeats of SREHP has enabled us to localize the native SREHP and demonstrate that the SREHP molecule could play a role in *E. histolytica* adhesion to mammalian cells. The ability to isolate *E. histolytica*-specific genes and to express these genes in *E. coli* may provide useful reagents for studying the molecular basis of *E. histolytica* pathogenesis and for future development of vaccines.

We thank J. Gordon, D. Krogstad, C. Rice, and E. Brown for their reading of this manuscript and many helpful suggestions, and C. Rice and T. Chambers for their assistance with the pATH expression system. E.L. is a Lucille P. Markey Scholar, and this work was supported in part by the Lucille P. Markey Charitable Trust Foundation.

1. Walsh, J. (1988) in *Amebiasis, Human Infection by Entamoeba histolytica*, ed. Ravdin, J. I. (Wiley, New York), pp. 93–105.
2. Ghadirian, E., Meerovitch, E. & Hartmann, D. P. (1980) *Am. J. Trop. Med. Hyg.* **29,** 779–784.
3. Krupp, I. M. (1974) *Am. J. Trop. Med. Hyg.* **23,** 355–360.
4. Swartzwelder, J. C. & Avant, W. H. (1952) *Am. J. Trop. Med. Hyg.* **1,** 567–575.
5. Li, E., Becker, A. & Stanley, S. L., Jr. (1989) *Infect. Immun.* **57,** 8–12.
6. Mattern, C. F. T. & Keister, D. B. (1977) *Am. J. Trop. Med. Hyg.* **27,** 882–887.
7. Diamond, L. (1968) *J. Parasitol.* **54,** 1047–1056.
8. Diamond, L. S., Harlow, D. R. & Cunnick, C. (1978) *Trans. R. Soc. Trop. Med. Hyg.* **4,** 431–432.
9. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18,** 5294–5299.
10. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
11. Gubler, U. & Hoffman, B. J. (1983) *Gene* **25,** 263–269.
12. Villa-Komaroff, L., Estratiadis, A., Broome, S., Lomedico, P., Tizard, R., Naker, S. P., Chick, W. L. & Gilbert, W. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 3727–3731.
13. Hanahan, D. (1983) *J. Mol. Biol.* **166,** 557–580.
14. Thomas, P. S. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 5201–5205.
15. Feinberg, A. P. & Vogelstein, B. (1983) *Anal. Biochem.* **132,** 6–13.
16. Li, E., Demmer, L. A., Sweetser, D. A., Ong, D. E. & Gordon, J. I. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 5779–5783.
17. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65,** 499–560.
18. Heuckeroth, R. O., Birkenmeier, E. H., Levin, M. S. & Gordon, J. I. (1987) *J. Biol. Chem.* **262,** 9709–9717.
19. Hardy, W. R. & Strauss, J. H. (1988) *J. Virol.* **62,** 998–1007.
20. Li, E., Becker, A. & Stanley, S. L., Jr. (1988) *J. Exp. Med.* **167,** 1725–1730.
21. Ozaki, L. S., Nussenzweig, R. S., Nussenzweig, V. & Godson, G. N. (1983) *Cell* **34,** 815–822.
22. von Heinje, G. (1986) *Nucleic Acids Res.* **14,** 4683–4690.
23. Tannich, E., Horstmann, R., Knobloch, J. & Arnold, H. H. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 5118–5122.
24. Kemp, D. J., Coppel, R. L. & Anders, R. F. (1987) *Annu. Rev. Microbiol.* **41,** 181–208.
25. Galinski, M. R., Arnot, D. E., Cochrane, A. H., Barnwell, J. W., Nussenzweig, R. S. & Enea, V. (1987) *Cell* **48,** 311–319.
26. Rodriguez, M. A., Hernandez, F., Santos, L., Valdez, A. & Orozco, E. (1989) *Mol. Biochem. Parasitol.* **37,** 87–100.